

Visión por Computadora 2024

Segundo Proyecto

07.marzo.2024

En este proyecto implementaremos algoritmos de Cascada para detección de objetos en imágenes y video. Haremos uso del algoritmo de Viola-Jones y de filtros de Haar.

Elegir alguna base de datos entre las siguientes:

- PASCAL VOC

<http://host.robots.ox.ac.uk/pascal/VOC/>.

(Yo sugiero usar la versión de 2012 para Action Classification Competition)

<http://host.robots.ox.ac.uk/pascal/VOC/voc2012/actionexamples/index.html>.

- Caltech Pedestrians Dataset

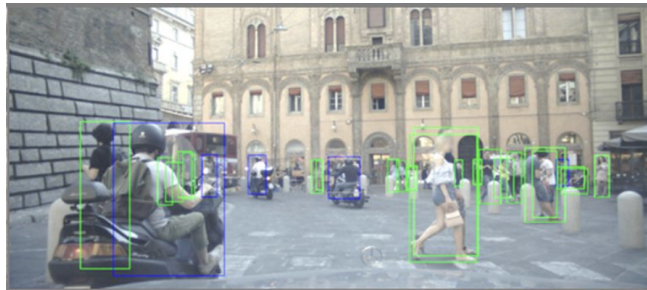
<https://data.caltech.edu/records/f6rph-90m20>.

- Eurocity Persons Dataset

<https://eurocity-dataset.tudelft.nl/>.

- PWR (Person Re-Identification in the Wild)

http://zheng-lab.cecs.anu.edu.au/Project/project_prw.html.



Problema 1: Construir un detector de personas (o peatones) en una imagen a color. Su detector debe mostrar un recuadro alrededor cada una de las personas localizadas en la imagen.

Elegir un subconjunto de 50 imágenes de alguno de los datasets anteriores, y mostrar ejemplo del desempeño de su detector. Como es habitual, se debe mostrar ejemplos donde el detector se desempeña correctamente y ejemplos donde falla.

Intentar explicar a partir de sus ejemplos, las limitaciones de este algoritmo.

Para el caso de detección vamos a definir:

- True positive (TP): si hay una persona en la imagen y el detector la encuentra.
- False positive (FP): si hay una ventana indicada por el detector, pero esta ventana no corresponde a una persona.
- False negative (FN): si hay una persona en la imagen, pero el detector no la encuentra.

Construir una tabla de métricas (*precision*, *recall*, *Jaccard index*) de clasificación binaria con las 50 imágenes elegidas.

Problema 2: Construir un detector de rostros para vídeo. Su detector debe mostrar un recuadro alrededor cada una de los rostros detectados, en cada uno de los *frames* donde son detectados.

Esta parte del proyecto se debe presentar en persona, y deberá poder usar una cámara (puede ser la cámara incluida en su dispositivo o alguna cámara *plug-in*). El objetivo de la presentación será evaluar el funcionamiento del detector en tiempo real.

Pueden usar alguna de las siguientes datasets para evaluar el funcionamiento de su detector de rostros.

- Labelled Faces in the Wild
<http://vis-www.cs.umass.edu/lfw/>.
- Fddb Face Detection Dataset and Benchmark
<http://vis-www.cs.umass.edu/fddb/index.html>.
- UTKFace
<https://susanqq.github.io/UTKFace/>.

Escribir un informe técnico con visualizaciones de los resultados obtenidos, la tabla de métricas y comparaciones, y sus principales conclusiones.

Evaluación : 10 puntos cada ejercicio, para un total de 20.

Fecha de Entrega: martes 5 de abril.